

# (Quasi-)linear time algorithm to compute LexDFS, LexUP and LexDown orderings

Arthur Milchior

## Abstract

We consider the three graph search algorithm LexDFS, LexUP and LexDOWN. We show that LexUP orderings can be computed in linear time by an algorithm similar to the one which compute LexBFS. Furthermore, LexDOWN orderings and LexDFS orderings can be computed in time  $(n + m \log m)$  where  $n$  is the number of vertices and  $m$  the number of edges.

## 1 Introduction

A graph search is a mechanism for systematically visiting the vertices of a graph. Deep-First Search (DFS) and Breadth-First Search (BFS) have been studied for decades (see e.g. [CSRL01]). Those two graph searches can be computed in linear time. A particular kind of BFS, the Lexicographical BFS (LexBFS), has then been introduced in [RTL76]. And by similarity, the Lexicographical DFS (LexDFS) has been studied in [CK08]. And then LexUP and LexDOWN in [Dus14].

A LexBFS ordering of a graph  $G$  is a possible output of a LexBFS search applied to  $G$ . While the LexBFS algorithm runs in time  $O(nm)$  where  $n$  is the number of vertices and  $m$  the number of edges, a LexBFS ordering can be computed in time  $O(n + m)$ . LexDFS, LexUP and LexDOWN also run in time  $O(nm)$ . We show that a LexUP ordering can be computed in linear time by an algorithm similar to the one which compute a LexBFS ordering. Furthermore, we prove that a LexDOWN ordering and a LexDFS ordering can be computed in time  $O(n + m \log m)$ .

Definitions are given in Section 2. The four graph search algorithms considered in this paper are given in Section 3. An efficient algorithm to compute LexDFS and LexDOWN ordering are given in Section 4. Finally, efficient algorithm is given to compute a LexUP ordering in Section 5.

## 2 Definition

Definitions used in this paper are now introduced. Most of those definitions are standard. Let  $\mathbb{N}$  be the set of non-negative integer. For  $A$  a finite set,  $|A|$  denotes the cardinality of  $A$ .

A word on  $\mathbb{N}$  is a sequence  $a_1 \dots a_n$ , with  $a_i \in \mathbb{N}$ . The empty word is denoted  $\epsilon$ . For  $\mathbf{a} = a_1 \dots a_n$  and  $\mathbf{b} = b_1 \dots b_m$  two words over  $A$ , it is said that  $\mathbf{a}$  is (lexicographically) smaller than  $\mathbf{b}$  if there exists  $i \leq \min(n, m)$  such that, for all  $1 \leq j \leq i$ ,  $a_j = b_j$ , and (either  $i = n < m$  or  $a_{i+1} < b_{i+1}$ ).

### 2.1 Graph

A (undirected) graph  $G$  with a source is a 3-tuple  $(V, E, s)$  where  $V$  is a finite set,  $E$  is a set of subsets of  $V$  whose elements's cardinality is 2 and  $s \in V$ . The elements of  $V$  are called vertices.

The elements of  $E$  are called edges. The vertex  $s$  is called the source.

A vertex  $v$  is said to be a neighbor of  $w$  if  $\{v, w\} \in E$ . The neighborhood of a vertex  $v$  is the set of neighbor of  $w$ , it is denoted  $N(v)$ . Formally,  $N(v) = \{w \mid \{v, w\} \in E\}$ . The degree of  $v$ , denoted  $d(v)$ , is the cardinality of its neighbourhood. Formally,  $d(v) = |N(v)|$ .

Two vertices  $v, w \in V$  are said to be connected if there exists a sequence  $v = v_0, \dots, v_p = w$  such that, for all  $0 \leq i < p$ ,  $\{v_i, v_{i+1}\} \in E$ . A graph is said to be connex if all pair of distinct vertices are connected.

## 2.2 Data structures

In this section, we list the data structures used in this paper. We list the operation those data structures admit, and their time complexity. All of those notions are standard (see e.g. [CSRL01]).

In this paper, each type is represented as *type*, each variable is represented as **var** and each function of parameter of an object  $o$  is represented as **o.param**.

It is assumed throughout this paper that *integers* can be incremented and compared in constant time. During execution of the algorithm of this paper of a graph  $(V, E)$ , all *integer* variables are interpreted by a number whose absolute value is at most  $\max(|V|, 2|E|)$ . Hence, the constant time assumption is relatively safe. Assigning a value  $x$  to a variable  $v$  is denoted  $v := x$  and is assumed to take constant time.

**Arrays** It is assumed in this paper that *arrays* are created in time linear to their numbers of elements. The elements of an *array*  $A$  with  $n$  elements are numbered from 1 to  $n$ . The  $i$ -th element of  $A$  is denoted  $A[i]$ , and can be read and assigned in constant time.

**Doubly linked lists** In this paper, all lists are assumed to be *doubly-linked lists*. A *doubly-linked list* of elements of type  $t$  is a sequence of *nodes*, with direct access to its first and last *nodes*. Each *node* contains a value of type  $t$ . Each *node* has also a direct access to its list, to the preceding and following *nodes*. A *doubly-linked list*  $l$  admits the following constant-time operations:

- Access to its first *node*:  $l.\mathbf{first}$ .
- Access to its last *node*:  $l.\mathbf{last}$ .
- Adding a *node*  $c$  to the head of  $l$ :  $l.\mathbf{add-first}(c)$ .
- Adding a *node*  $c$  to the end of  $l$ :  $l.\mathbf{add-last}(c)$ .

A list with  $n$  *nodes* can be sorted in time  $O(n \cdot \log(n))$ , assuming that the comparison of two *nodes* of the list can be done in constant time:  $l.\mathbf{sort}$ . The order will always be clear in the algorithms of this paper.

A *node*  $e$  of a *doubly-linked list*  $l$  admits the following operations:

- access to the preceding *node*:  $e.\mathbf{pred}$ ,
- access to the following *node*:  $e.\mathbf{next}$ ,
- access to the value at position  $e$ :  $e.\mathbf{value}$ ,
- inserting a value  $i$  of type  $t$  in a new *node* after  $e$ :  $e.\mathbf{add-after}(i)$  and

- inserting a value  $i$  of type  $t$  in a new *node* before  $e$ :  $e.\text{add-before}(i)$  and
- removing  $e$ :  $e.\text{remove}$ .

Note that the first value of type  $t$  of a *list*  $l$  is  $l.\text{first.value}$  and not  $l.\text{first}$ . Indeed,  $l.\text{first}$  is a *node* and not a value of type  $t$ .

**Graphs** A *graph*  $G$  is represented as an *array* of size  $n$ . The  $i$ -th element of the *array* contains the list of neighbors of  $v_i$ . Formally,  $N(i)$  should be represented as  $G[i]$ , however,  $N(i)$  is used in the algorithms of this paper for the sake of the readability.

### 3 Graph search algorithm

In this section, the four graph search algorithms considered in this paper are considered. A graph search algorithm is an algorithm as in Algorithm 1. Note that the standard definition of graph search algorithms is more general than the one used in this paper. The only difference between the

<b>Algorithm 1:</b> Definition of a graph search algorithm	
<b>Input:</b> An undirected <i>graph</i> $G = (V, E, s)$ with $n$ vertices	
<b>Output:</b> an ordering $\sigma$ of the vertices of $G$	
2	assign $\text{.label}$ $\epsilon$ to all vertices;
4	assign $\text{.label} [\infty]$ to $s$ ;
6	<b>foreach</b> $i$ from 1 to $n$ <b>do</b>
8	pic an <b>unnumbered vertex</b> $\text{vertex}$ with lexicographically maximal $\text{.label}$ ;
10	$\sigma[i] := \text{vertex}$ ;
12	<b>foreach</b> <i>unnumbered vertex</i> $\text{neighb} \in N(\text{vertex})$ <b>do</b>
14	update $\text{neighb.label}$ ;
15	Output $\sigma$ ;

four graph search algorithms considered in this paper appears in Line 12 of Algorithm 1.

In this paper, the label is always a *list* of *integers*. Each update always takes constant time and add exactly an *integer* to the label. The time complexity of this algorithm is now considered.

**Lemma 3.1.** *Let  $G = (V, E, s)$  a graph with  $n$  vertices and  $m$  edges. Assuming the update consists in adding an integer in the front or in the rear of the list, the time complexity of Algorithm 1 is  $O(nm)$ .*

*Proof.* Let us first consider the labels. The label of a vertex  $v$  contains at most  $|N(v)|$  elements. Hence the sum of the length of the label is at most  $2m$ .

Lines 2, 4 are executed once and in constant time. Hence their time cost is  $O(1)$ . Each execution of Line 10 may have to read the entire labels of each vertex. Finding the maximal label then cost  $O(m)$ -times. Since this line is executed  $n$  times, this Line costs  $O(nm)$ -times. Line 10 can be executed in constant time, and is executed  $n$  times, hence it costs  $O(n)$  time. Finally, each execution of line 14 takes constant time. And this line is executed once for each  $1 \leq i \leq n$  and  $n \in N(v_i)$ , hence it is executed  $2m$ -times. Thus, it costs  $O(m)$  time.

Finally, the whole algorithm runs in time  $O(nm)$ . □

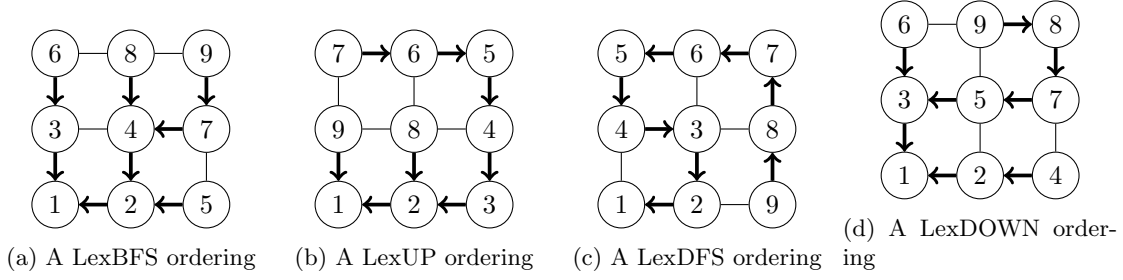


Figure 1: Orderings with source in a corner

For  $\mathcal{A}$  a graph search algorithm, an  $\mathcal{A}$  ordering of  $G$  is a possible output of  $\mathcal{A}$  on  $G$ . Those four algorithms are now defined, as in [Dus14].

### 3.1 Lexicographic Breadth-First Search

Let Lexicographic Breadth-First Search (LexBFS) be a graph search algorithm, as in Algorithm 1, where Line 12 is: “append  $n - i$  to the **neighb**’s label”.

Intuitively, at each step, the vertex  $v$  is preferred to the vertex  $v'$  if the first numbered neighbor of  $v$  have been numbered earlier than the first numbered neighbor of  $v'$ . If their first neighbor are equal, then the same comparison is done on the second neighbor. And so on. If  $v$  and  $v'$  have  $i$  and  $i'$  numbered neighbors respectively, with  $i' < i$ , and furthermore if the  $i'$  first numbered neighbors of  $v$  are exactly the first  $i$  numbered neighbors of  $v'$  in the same order, then  $v$  is also preferred.

Figure 1a shows examples of LexBFS ordering. Each arrow associates to a vertex  $v$  its earliest numbered neighbor. Table 1 associate to each vertex  $v$  its list of numbered neighbors when  $v$  was numbered. This table also associate to  $v$  its label.

Vertex number	2	3	4	5	6	7	8	9
$v$ ’s label in Figure 1a	8	8	76	7	6	54	53	21
When $v$ is numbered, it’s numbered neighbors are	1	1	23	2	3	45	46	78

Table 1: Label during the LexBFS search of Figure 1a

### 3.2 Lexicographic UP

Let Lexicographic UP (LexUp) be a graph search algorithm, as in Algorithm 1, where Line 12 is: “append  $i$  to the **neighb**’s label”.

Intuitively, at each step, the vertex  $v$  is preferred to the vertex  $v'$  if the first numbered neighbor of  $v$  have been numbered later than the first numbered neighbor of  $v'$ . If their first neighbor are equal, then the same comparison is done on the second vertex. And so on. If  $v$  and  $v'$  have  $i$  and  $i'$  numbered neighbors respectively, with  $i' < i$ , and furthermore if the  $i'$  first numbered neighbors of  $v$  are exactly the first  $i$  numbered neighbors of  $v'$  in the same order, then  $v$  is also preferred.

Note that this intuition is the same than for LexBFS, apart that the word “earlier” have been replaced by the word “later”. It is because, in both cases, *integers* are prepended to the label. But in

the former case, the sequence of prepended numbers decrease while in the second case it increases. Thus, the maximal numbers are added earlier in LexBFS and later in LexUP.

Figure 1b show an example of a LexBFS ordering. Each arrow associates to a vertex  $v$  its first numbered neighbor. Table 2 associates to each vertex  $v$  its label when  $v$  was numbered in each of those 4 examples respectively. Note that this list is also its list of numbered neighbors.

Vertex number	2	3	4	5	6	7	8	9
$v$ 's label in Figure 1b	1	2	3	4	5	6	246	187

Table 2: Label during the LexUP search of Figure 1

### 3.3 Lexicographic Depth-First Search

Let Lexicographic Depth-First Search (LexDFS) be a graph search algorithm, as in Algorithm 1, where Line 12 is: “prepend  $i$  to the **neighb**'s label”.

Intuitively, at each step, the vertex  $v$  is preferred to the vertex  $v'$  if the last numbered neighbor of  $v$  have been numbered later than the last numbered neighbor of  $v'$ . If their last neighbor are equal, then the same comparaisn is done on the second last neighbor. And so on. If  $v$  and  $v'$  have  $i$  and  $i'$  numbered neighbors respectively, with  $i' < i$ , and furthermore if the  $i'$  first numbered neighbors of  $v$  are exactly the last  $i$  numbered neighbors of  $v'$  in the same order, then  $v$  is also preferred.

Note that this intuition is the same than for LexUP, apart that the word first have been replaced by the word last. Indeed the same integers is added to the label in both cases. However, in the former case the integer is prepended while in the latter case the integer is appended. Hence, in both cases, neighbors with small number are preferred. But in LexUP they must be the earliest neighbors while in LexDFS they must be the latest neighbors.

Figure 1c show an example of a LexBFS ordering. Each arrow associates to a vertex  $v$  its last numbered neighbor. Table 3 associates to each vertex  $v$  its label when  $v$  was numbered. Note that this list is also its list of numbered neighbors.

Vertex number	2	3	4	5	6	7	8	9
$v$ 's label in Figure 1c	1	2	31	4	53	6	73	82

Table 3: Label during the LexDFS search

### 3.4 LexDown

Let Lexicographic DOWN be a graph search algorithm, as in Algorithm 1, where Line 12 is: “prepend  $n - i$  to the label of **neighb**”.

Intuitively, at each step, the vertex  $v$  is preferred to the vertex  $v'$  if the first numbered neighbor of  $v$  have been numbered later than the first numbered neighbor of  $v'$ . If their first neighbor are equal, then the same comparaisn is done on the second neighbor. And so on. If  $v$  and  $v'$  have  $i$  and  $i'$  numbered neighbors respectively, with  $i' < i$ , and furthermore if the  $i'$  first numbered neighbors of  $v$  are exactly the last  $i$  numbered neighbors of  $v'$  in the same order, then  $v$  is also preferred.

Note that this intuition is the same than for LexBFS (respectively, LexDFS), apart that the word earlier (respectively, first) have been replaced by the word later (respectively, last). The reason is similar to the previous explanations.

Figure 1d shows an examples of LexBFS ordering. Each arrow associates to a vertex  $v$  its last numbered neighbor. Table 4 associates to each vertex  $v$  its list of numbered neighbors when  $v$  was numbered. This table also associate to  $v$  its label. Note that when vertices 1 and 2 are fixed, this

Vertex number	2	3	4	5	6	7	8	9
$v$ 's label in Figure 1d	8	8	7	67	6	45	2	134
When $v$ is numbered, its numbered neighbors are	1	1	2	32	3	54	7	865

Table 4: Labels during the LexDOWN search

graph admits no other LexBFS ordering.

## 4 LexDFS and LexDOWN

An algorithm is now given, which outputs a LexDOWN ordering in time  $O(n + m \log(m))$ . Note that  $m \leq n^2$ , hence  $\log(m) \leq 2 \log n$ , thus, this algorithm is more efficient than Algorithm 1.

**Theorem 4.1.** *Let  $G = (V, E, s)$  be a connex undirected graph with source  $s$ , with  $n$  vertices and  $m$  edges. A LexDFS ordering of  $G$  can be computed in time  $O(n + m \log(m))$ .*

An intuition of the algorithm is first given. Note that, in Algorithm 1, at each iteration of the loop of Line 10, all unnumbered states must be checked. At each iteration, this line runs in time  $O(m)$ . This time can be avoided if the list is already sorted. Since at the  $i$ -th iteration, at most  $|N(v_{\sigma(i)})|$  labels change, it suffices to sort and move those  $O(|N(v_{\sigma(i)})|)$  elements. The sorting can be done in time  $O(|N(v_{\sigma(i)})| \log(|N(v_{\sigma(i)})|))$ . Since all of those elements must be moved to the front of the *list*, a correct usage of pointers allow to move the  $O(|N(v_{\sigma(i)})|)$  vertices in time  $O(|N(v_{\sigma(i)})|)$ . Summing over all  $i$ , the times taken by those operations is  $O(\sum_{i=1}^n |N(v_{\sigma(i)})| \log(|N(v_{\sigma(i)})|)) = m \log(m)$ .

A simplified version of the algorithm is given as Algorithm 2. In this simplified version, *vertex* is a type which contains an *integer order* and a **label**. Algorithm 3 furthermore shows exactly how to use pointers in order to obtain a quasi-linear time.

*Proof.* In Algorithm 3, a *vertex* is a data structure which contains 4 parameters

- **order** : an *integer*;
- **pos** : a *node* of a *list* of *integers*;
- **label** : a *list* of *integers*;
- **numbered** : a *Boolean*;

Let us first prove that Algorithm 3 returns a lexDFS ordering. For  $1 \leq i \leq n$ , let  $\sigma_i$ , **unnumbered** <sub>$j$</sub> , vertices <sub>$j$</sub>  and **max** <sub>$j$</sub>  be the values of those variables when the iteration of the loop of Line 16 ends, with the variable  $i$  interpreted by  $j$ . Finally, let  $\sigma_0$ , **unnumbered**<sub>0</sub>, vertices<sub>0</sub> and **max**<sub>0</sub> be the values of those variables before the first iteration of this loop.

The loop invariants of this algorithm are:

**Algorithm 2:** Computing a LexDFS ordering-simplified

**Input:**  $G = (V, E, s)$  an undirected *graph* with a source  
**Output:** a LexDFS-Simple ordering  $\sigma$  of the vertices of  $G$

```

2  $\sigma$ : array of  $n$  integers;
4 vertices: array of  $n$  elements of type vertex;
6  $\text{max} := 0$ ;
8 foreach  $i$  from 1 to  $n$  do                                     /* Initialization */
10 |   vertices[ $i$ ] := {order :=  $-\infty$  label := []};
12 vertices[ $s$ ] := {order := 0; label := [ $\infty$ ]};
14 unnumbered := [ $s$ ];
16 foreach  $i$  from 1 to  $n$  do
18 |    $\sigma(i) := \text{unnumbered.first.value}$ ;                       /* Selecting the greatest value. */
20 |   remove  $\sigma(i)$  from unnumbered;
22 |   sort the neighbors of  $v_{\sigma(i)}$  in increasing order;
24 |   foreach neighb: unnumbered neighbor of  $v_{\sigma(i)}$  in increasing order do
26 |   |   if neighb's label is empty then                         /* neighb must be removed from */
28 |   |   |   remove neighb from unnumbered;                     /* unnumbered if its was in it. */
30 |   |   prepend  $i$  to neighb's label;                          /* neighb now has the greatest label */
32 |   |   add neighb to the front of unnumbered;
34 |   |   set  $\text{max}$  to  $\text{max} + 1$ ;
36 |   |   set neighb's order to  $\text{max}$ ;                             /* and has the greatest order */
37 return  $\sigma$ 

```

1.  $\sigma_j[i]$  contains an element  $k$  such that  $\text{label}_i(v_k)$  is lexicographically maximal, for  $0 < i \leq j$ .
2.  $\text{vertices}_j[i].\text{label}$  contains the label of  $v_i$ , as in the  $j$ -th step of LexDFS. Note that  $\text{vertices}[i].\text{label}$  is not actually used in computation of the LexDFS ordering.
3. The variable  $\text{unnumbered}_j$  contains the *list* of **unnumbered** vertices with a non-empty label. Those vertices appears in decreasing lexicographic order of their labels.
4. If  $x$  appears before  $y$  in  $\text{unnumbered}_j$ , then  $\text{node}_j[x].\text{order} > \text{node}_j[y].\text{order}$ .
5. If  $\text{unnumbered}_j$  contains the vertex  $v_i$ , then  $\text{vertices}_j[i].\text{pos}$  is the *node* of  $\text{unnumbered}_j$  whose value is  $v_i$ . Otherwise,  $\text{vertices}_j[i].\text{pos}$  is unspecified.
6.  $\text{max}_j$  is greater than all finite  $\text{vertices}_j[i].\text{order}$ .
7.  $\text{max}_j$  is less than the sum of the degree of the vertices  $v$  which are numbered at the  $j$ -th step.

Let us show that, for  $0 \leq j \leq n$ , the 7 invariant are satisfied. Invariant 5 is satisfied at each step, because everytime an *integer*  $i$  is added into **unnumbered**,  $\text{vertices}[i].\text{pos}$  is modified accordingly. The proof for the other invariants is by induction on  $j$ .

**Algorithm 3:** Computing a LexDFS ordering

**Input:**  $G = (V, E, s)$  an undirected *graph* with a source  
**Output:** a LexDFS ordering  $\sigma$  of the vertices of  $G$

```

2  $\sigma$ : array of  $n$  integers initialized to  $-1$ ;
4 vertices: array of  $n$  elements of type vertex;
6  $\text{max} := 0$ ;
8 foreach  $i$  from 1 to  $n$  do                                /* Initialization */
10 |   vertices[ $i$ ] := {order :=  $-\infty$ ; label := []; numbered := false};
12 unnumbered := [ $s$ ];
14 vertices[ $s$ ] := {order := 0; pos := unnumbered.last; label := [ $\infty$ ]; numbered := false};
16 foreach  $i$  from 1 to  $n$  do
18 |    $\sigma(i) := \text{unnumbered.first.value}$ ;                    /* Selecting the greatest value. */
20 |   unnumbered.first.remove;
22 |   vertices[ $\sigma(i)$ ].numbered := true;
24 |   sort the neighbors of  $v_{\sigma(i)}$  in increasing order;
26 |   foreach neighb: unnumbered neighbor of  $v_{\sigma(i)}$  in increasing order do
28 |   |   if vertices[neighb].label  $\neq []$  then                /* neighb must be removed from */
30 |   |   |   vertices[neighb].pos.remove;                    /* unnumbered if its was in it. */
32 |   |   vertices[neighb].label.add-first( $i$ ); /* neighb now has the greatest label */
34 |   |   unnumbered.add-first(neighb); /* hence, it goes in front of the list, */
36 |   |   vertices[neighb].pos := unnumbered.first;
38 |   |    $\text{max} := \text{max} + 1$ ;
40 |   |   vertices[neighb].order :=  $\text{max}$ ;                    /* and has the greatest order */
41 return  $\sigma$ 

```

Let us show that, for  $j = 0$ , the 7 invariant are satisfied.

Invariant 1 holds, since there are no integer  $0 < i \leq 0$ .

By definition of LexDFS, all labels are empty at initialization, apart from the one of the source. It is the case in this program because of Lines 10 and 14. Hence invariant 2 is satisfied.

Note that  $v_s$  is the only labelled vertex and that no vertex is numbered. Furthermore  $s$  is the element of **unnumbered** because of Line 12. Hence invariant 3 is satisfied.

Invariant 4 is also trivially satisfied, since  $s$  have the greatest order and the greatest label, and all other orders are equal and all other label are equals.

Invariant 6 is trivially satisfied since for all  $i$ , vertices<sub>0</sub>[ $i$ ].**order** = 0.

Invariant 7 is trivially satisfied since no vertices are numbered at the 0-th step.

Let  $0 < j \leq n$ . Let us now assume that the 7 invariants holds at step  $j - 1$ , and let us prove that it holds for  $j$ .

Since Invariant 1 holds at step  $j - 1$ , it clearly holds at step  $j$  for all  $i < j$ . It remains to consider the case  $i = j$ . By invariant 3, **unnumbered** <sub>$j$</sub>  contains the *list* of **unnumbered** labelled vertices at step  $j$ , in decreasing lexicographic order of their labels. Hence Line 12 correctly assigns to  $\sigma[j]$  a vertex  $w$  such that label <sub>$j$</sub> ( $w$ ) has a maximal label. Thus, Invariant 1 holds at step  $j$ .

Invariant 2 clearly remains true since the updating of the label is exactly the one of the definition of the LexDFS algorithm.



At the  $j$ -th step, the list of **unnumbered** vertices with a non-empty label contains, in this order:

- The neighbors of  $v_{\sigma(j)}$ , which are **unnumbered** and have a non-empty label at step  $j - 1$ . The order, according to their labels, are in the same order in both lists.
- The vertices which are neither  $v_{\sigma(j)}$  nor its neighbors, which are **unnumbered** and have a non-empty label at step  $j - 1$ . The order, according to their labels, are in the same order in both *lists*.
- The neighbors of  $v_{\sigma(j)}$  which are **unnumbered** and have an empty label at step  $j - 1$ .

Thus, according to invariant 3, **unnumbered<sub>j</sub>** must contains, in the following order:

- the elements of **unnumbered<sub>j-1</sub>** which are neighbors of  $v_{\sigma(j)}$ , in the same order,
- the elements of **unnumbered<sub>j-1</sub>** which are neither neighbors of  $v_{\sigma(j)}$  nor  $j$ , in the same order,
- the unlabelled neighbors of  $v_{\sigma(j)}$ , in an arbitrary order.

This is indeed the value of **unnumbered<sub>j</sub>**, because of Lines 30 and 34. Hence invariant 3 holds at step  $j$ .

Since each time an element is moved to the front of **unnumbered**, its **order** is greater than any **order** presently assigned 5, its **order** is greater than any previously assigned **order**, then Invariant 4 holds.

Invariant 6 clearly holds since the **orders** are assigned in increasing order, and since, each time an **order** is assigned, **max** is assigned to be its predecessor.

It is easy to see that  $\mathbf{max}_j \leq \mathbf{max}_{j-1} + |N(\sigma(j))|$ . Hence invariant 7 is true at step  $j$ .

Since the invariants are satisfied at each steps, by 1, at the end of the loop,  $\sigma$  contains a LexDFS ordering of  $G$ . Hence the algorithm indeed returns a LexDFS ordering of  $G$ .

Let us now consider the computation time. The code of Lines 2, 4, 6, 14 and 12 are executed exactly once, and runs in time  $O(n)$ . Hence their cost is  $O(n)$ .

Lines 10, 18, 20, 22, are executed  $n$  times and runs in constant time. Hence their cost is  $O(n)$ .

Line 24 is executed once for each vertex  $v_i$ . And for each vertex  $v_i$ , it runs in time  $O(|N(v_i)| \log(|N(v_i)|))$ . Hence the total cost of this line is  $O(\sum_{i=1}^n |N(v_i)| \log(|N(v_i)|)) = O(m \log m)$ .

Lines 32 to 40 are executed once by edge, and executed in constant time. Hence their cost is  $O(m)$ .

Finally, the total execution time is  $O(n + m \log(m))$ .  $\square$

Note that the **orders** are either infinite, or *integers* between 0 and  $2m$ . Hence it is acceptable to assume that comparison of two **order** parameters can be done in constant time.

**LexDOWN** As stated in Section 3.4, LexDOWN is similar to LexDFS. It is now considered.

**Theorem 4.2.** *Let  $G = (V, E, s)$  be a connex undirected graph with source  $s$ , with  $n$  vertices and  $m$  edges. A LexDOWN ordering of  $G$  can be computed in time  $O(n + m \log(m))$ .*

*Proof.* The algorithm to compute a LexDOWN ordering is Algorithm 3, with the three following changes:

- Line 32 is tranformed into “vertices[neighb].label.prepend.( $n - i$ )”,

- Line 34 is transformed into “**unnumbered.add-last.(neighb);**” and
- Line 38 is transformed into “**max:=max-1;**”.

Invariant 6 must be changed to “**max<sub>j</sub>** is smaller than all finite vertices<sub>j</sub>[*i*].**order**”, and 7 must be changed to “|**max<sub>j</sub>**| is less than the sum of the degree of the vertices *v* which are numbered at the *j*-th step”. Apart from those changes, the proof of this theorem is exactly the same than the proof of Theorem 4.1.  $\square$

## 5 Efficient LexBFS and LexUP

In this section, it is shown that a LexUP ordering can be computed in linear time. The algorithm is very similar to the algorithm for efficiently computing LexBFS.

A simplified version of the linear time algorithm which computes a LexBFS ordering is recalled as Algorithm 4. This algorithm keeps a *list*, **unnumbered**, which contains all vertices, with a non-empty label, in decreasing order according to their label. More precisely, all (indices of) vertices with the same non-empty label belong to a *set*, and **unnumbered** is a *list* of *sets*. The *sets* are also encoded as *lists*. When a vertex  $v_i$  is numbered, the label of its neighbors increases. However, it does not increase enough to become greater than labels which used to be greater than it. Hence all neighbors belonging to the same *set* *s* are moved to a new *set* *s'* placed before *s*. As soon as a *set* is empty, it is removed from the list. Each vertex *v* is moved at most  $|N(v)|$  times in the list.

A correct usage of pointers, as shown in Algorithm 5, allows to move indices from the previous *set* to the new *set* in constant time. Hence, the algorithm runs in time  $O(n + m)$ . In this algorithm, a *set* is a data-structure with three parameters:

- **pos**: a *node* of a *list* of *sets*,
- **edited**: an *integer* and
- **elements**: a *list* of *integers*.

And a *vertex* is a data-structure with four parameters:

- **pos**: a *node* of a *list* of *integers*,
- **numbered**: a *Boolean*,
- **label**: a *list* of *integers* and
- **set**: a *set*.

Note that if a vertex *v* have an empty label, it has the lexicographically smallest label. Hence, when a first element is added to the label of *v*, this vertex moves to the second least set (which may become the least set if there remains no more vertex with an empty label). Indeed, the first element of *v*'s label is  $n - i$ . And  $n - i$  is smaller than the first element of the label of all other vertices *w* with non-empty label.

The preceding remark leads to the main difference between LexBFS and LexUP. In LexUP, the element added is *i*, and not  $n - i$ . Hence the first element of this vertex *v* is *i*. Hence, it is greatest than all the first element of all other vertices *w* with a non-empty label. Hence, *v* moves to the greatest set. Therefore, to transform Algorithm 5 into an algorithm which computes a LexUP ordering, it suffices to do the following change:

**Algorithm 4:** Efficient computation of a LexBFS - simplified

**Input:** An undirected *graph*  $G = (V, E)$   
**Output:** an ordering  $\sigma$  of the vertices of  $G$

```

1 vertices: array of  $n$  elements of type vertex;
2 foreach  $i$  from 1 to  $n$ , distinct from  $s$  do                                /* Initialization */
3   |  $v_i$ 's label is set to  $[]$ ;
4   set_ $s$  is set to  $[s]$ ;
5    $s$ 's label is set to  $[\infty]$ ;
6   unnumbered :=  $[set\_s]$ ;
7   foreach  $i$  from 1 to  $n$  do
8     | greatest_set := the first element of unnumbered;
9     |  $\sigma[i]$  := any element of greatest_set;                                /* Selecting a greatest vertex */
10    | Remove this element from greatest_set; /* and removing it from the list. */
11    | If greatest_set is empty, remove it from unnumbered;
12    | foreach  $neighb \in N(v_{\sigma(i)})$ , unnumbered do
13      | if  $neighb$ 's label is not empty then
14        | if no vertices from  $neighb$ 's set have been seen for this value of  $i$  then
15          | new_set is set to the  $[]$ ;
16          | add new_set before  $neighb$ 's set;
17        | else
18          | set new_set to the set preceding  $neighb$ 's set;
19          | If  $neighb$ 's set is a singleton, remove this set from unnumbered;
20        | else                                /* If  $neighb$ 's label is empty */
21          | if No unlabelled neighbor have been seen for this value of  $i$  then
22            | Set new_set to a new set;
23            | Add new_set to the rear of unnumbered;
24            | Set new_set to the last set of unnumbered;
25          | Move  $neighb$  to new_set;
26          | append  $n - i$  to  $neighb$ 's label
27  return  $\sigma$ ;

```

- Line 36 must be modified to “**unnumbered.add-first(new\_set);**”.
- Line 44 must be changed to “**vertices[neighb].label.add-last(i);**”.

The proof that Algorithm 3 computes a LexDFS ordering is similar to the proof that Algorithm 5 computes a LexBFS ordering.

Note that, if **unlabelled** was not restricted to contains only labelled vertices, the algorithm would still be correct for LexBFS. Furthermore, the algorithm would be shorter. However, the algorithm will not be correct anymore for LexUP.

## 6 Conclusion

In this paper, it has been proven that a LexUP ordering can be computed in linear time and that a LexDOWN ordering and a LexDFS ordering can be computed in time  $O(n + m \log m)$ .

The author thanks Michel Habib, who introduced this problem to him during his Graph Theory Lectures.

## References

- [CK08] Derek G. Corneil and Richard Krueger. A unified view of graph searching. *SIAM J. Discrete Math.*, 22(4):1259–1276, 2008.
- [CSRL01] Thomas H. Cormen, Clifford Stein, Ronald L. Rivest, and Charles E. Leiserson. *Introduction to Algorithms*. McGraw-Hill Higher Education, 2nd edition, 2001.
- [Dus14] Jérémie Dusart. *Graph Searches with Applications to Cocomparability Graphs*. Theses, Université Denis Diderot Paris 7, June 2014.
- [RTL76] Donald J. Rose, Robert Endre Tarjan, and George S. Lueker. Algorithmic aspects of vertex elimination on graphs. *SIAM J. Comput.*, 5(2):266–283, 1976.

**Algorithm 5:** Efficient computation of a LexUP

**Input:** An undirected *graph*  $G = (V, E)$   
**Output:** an ordering  $\sigma$  of the vertices of  $G$

```

1 vertices: array of  $n$  elements of type vertex;
2 unlabelled_edited:=0;
3
4 foreach  $i$  from 1 to  $n$ , distinct from  $s$  do                                /* Initialization */
5   | vertices[ $i$ ]:={numbered:=false, label:=[]};
6   set_s:={pos:=set_s;edited:=0;elements:=[ $s$ ]};
7   vertices[ $s$ ]:={pos:=set_s.elements.first,numbered:=false, label:=[ $\infty$ ];set:=set_s};
8   unnumbered:=[set_s];
9   foreach  $i$  from 1 to  $n$  do
10    | greatest_set:=unnumbered.first.value;
11    |  $\sigma[i]$  :=greatest_set.elements.first.value;      /* Selecting a greatest vertex */
12    | greatest_set.elements.first.remove;      /* and removing it from the list. */
13    | vertices[ $\sigma(i)$ ].numbered:=true;
14    | if  $\text{greatest\_set.elements}=[]$  then
15    |   | greatest_set.pos.remove;
16    |   foreach  $neighb \in N(v_{\sigma(i)})$ , unnumbered do
17    |     | if  $\text{vertices}[neighb].\text{label} \neq []$  then /* If the neighbor's label is not empty */
18    |       | if  $\text{vertices}[neighb].\text{set.edited} < i$  then /* no neighbors with the same label
19    |         | have been seen: a new set must be created before the current one.
20    |         | */
21    |         | vertices[neighb].set.edited:= $i$ ;
22    |         | new_set:={edited:= $i$ ; elements:=[]};
23    |         | vertices[neighb].set.pos.add-before(new_set);
24    |         | new_set.pos:= vertices[neighb].set.prec;
25    |       | else /* A neighbor with the same label have already been seen */
26    |         | new_set:=vertices[neighb].set.pos.prec;
27    |         | vertices[neighb].pos.remove;
28    |         | if  $\text{vertices}[neighb].\text{set.elements}=[]$  then
29    |         |   | vertices[neighb].set.remove;
30    |         | else
31    |         |   /* If neighb's label is empty */
32    |         |   if  $\text{unlabelled\_edited} < i$  then /* No unlabelled neighbors have been
33    |         |     | considered yet. */
34    |         |     | unlabelled_edited:= $i$ ;
35    |         |     | new_set:={edited:= $i$ ; elements:=[]};
36    |         |     | unnumbered.add-last(new_set);
37    |         |     | new_set.pos:= unnumbered.last;
38    |         |   else
39    |         |     new_set:= unnumbered.last.value;
40    |         | new_set.elements.add-last(neighb); /* Moving the neighbor */
41    |         | vertices[neighb].set:=new_set;
42    |         | vertices[neighb].pos:=new_set.elements.last;
43    |         | vertices[neighb].label.add-last( $n - i$ ); /* Updating the label of neighb */
44   return  $\sigma$ ;

```

## Index

$(V, E, s)$ , 1  
(undirected) graph with a source, 1  
 $|A|$ , 1  
 $\mathcal{A}$  ordering, 4  
Cardinality, 1  
Connected vertices, 2  
Connex graph, 2  
 $d(v)$ , 2  
Edges, 2  
Empty word, 1  
 $\epsilon$ , 1  
LexBFS, 4, 5  
LexDOWN, 5  
Lexicographic Breadth-First Search, 4  
Lexicographic Depth-First Search, 5  
Lexicographic DOWN, 5  
Lexicographic UP, 4  
Lexicographically smaller, 1  
LexUP, 4  
 $N(v)$ , 2  
 $\mathbb{N}$ , 1  
neighbor, 2  
Neighborhood of a vertex, 2  
Source of a graph, 2  
Vertices, 1  
Word, 1